

BIREDIAL 2013
Conferencia Internacional Acceso Abierto, Preservación Digital y Datos Científicos
San José (Costa Rica), 15-17 de octubre de 2013

Preservación digital para instituciones de tamaño medio: una adaptación pendiente

Miquel Térmens Graells

Universidad de Barcelona

Departamento de Biblioteconomía y Documentación

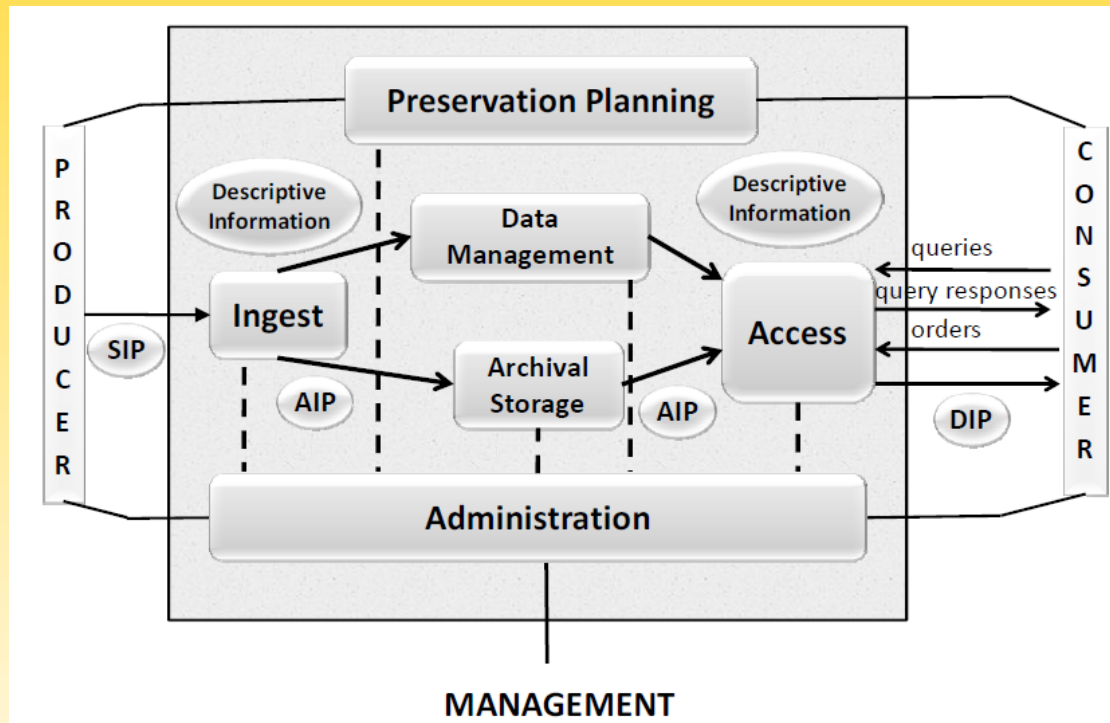
termens@ub.edu

Programa

1. El problema
2. La investigación en curso
3. Resultados
4. Conclusiones

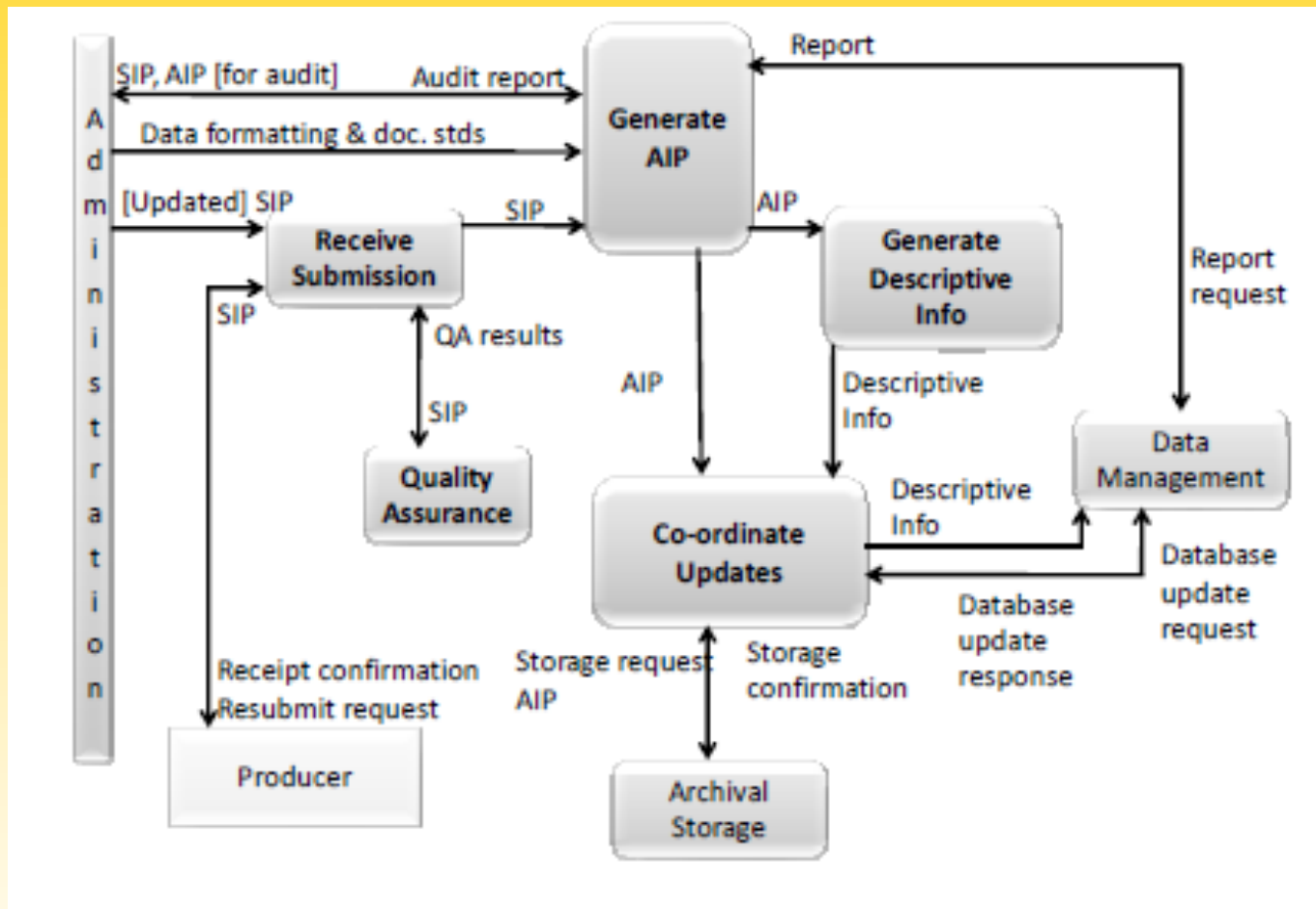
1. El problema

OAIS es complicado



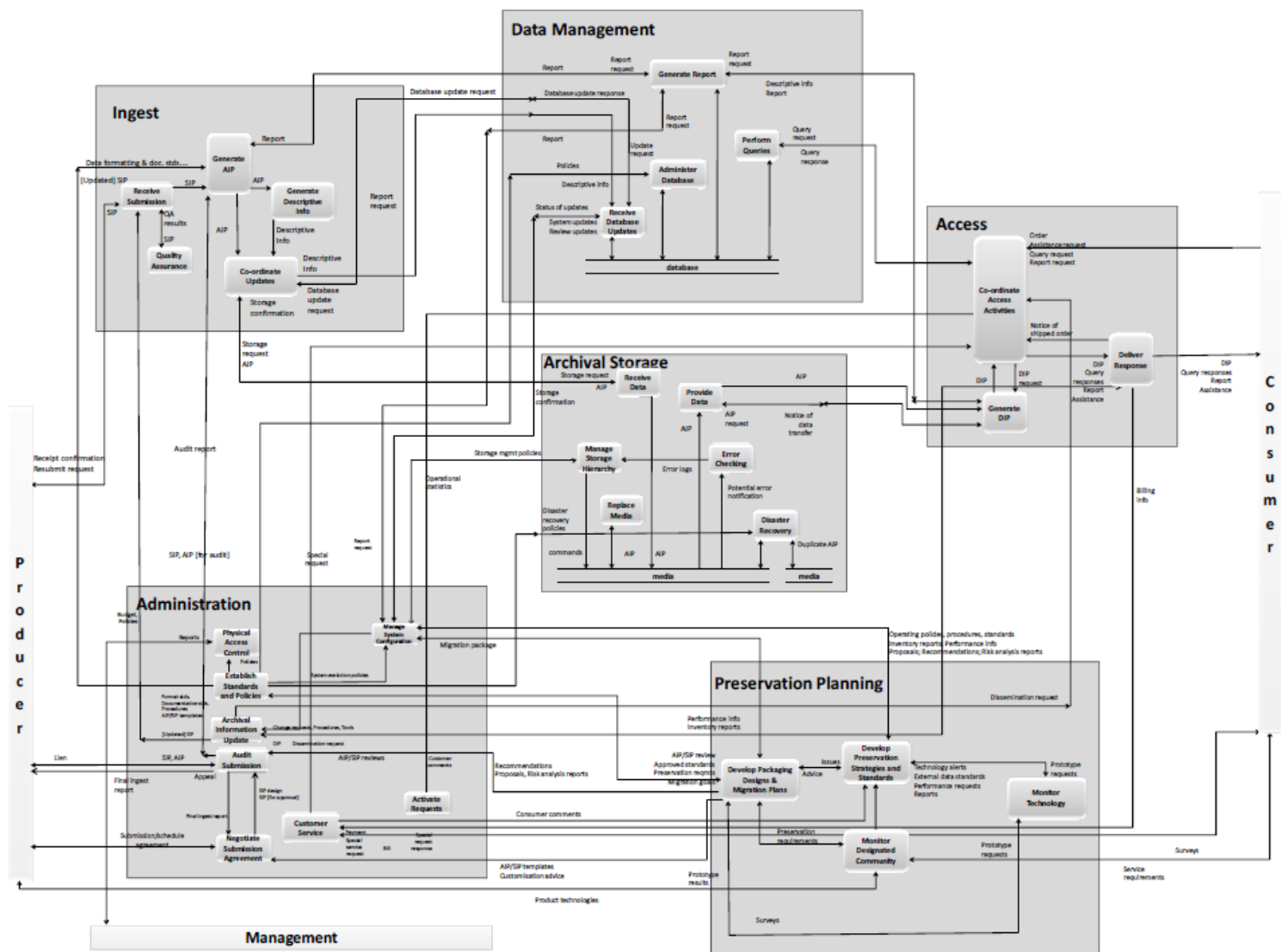
Fuente: OAIS, June 2012, page 4-1

OAIS es más complicado



Fuente: OAIS, June 2012, page 4-5

OAIS es extremadamente complicado



Y si esto solamente es un esquema...

¿Qué habrá por debajo?...



Apreciaciones observadas

- OAIS es complicado.
- OAIS solo lo pueden comprender los expertos.
- Preservar es técnicamente complicado.
- Preservar es caro.
- No disponemos de la formación para dirigir o gestionar un proyecto de este tipo.
- No sabemos como “vender” este tema a las autoridades.
- La preservación solo está al alcance de las grandes instituciones internacionales.



No preservamos

2. La investigación en curso

Investigaciones

- 2 Auditorías de seguridad informática y de preservación digital en 8 hospitales de Barcelona (ENS y TDR ISO/IEC 16363:2012).
- Auditoría de preservación digital en el Instituto Cartográfico de Cataluña (ISO/IEC 16363:2012).
- 2 Auditorías de seguridad informática y de preservación digital en la Universidad Autónoma de Barcelona (ENS y TDR ISO/IEC 16363:2012).
- Estudio de creación de una unidad de análisis forense digital en una biblioteca nacional.
- Estudio de los formatos preservados en los repositorios institucionales UB, RECERCAT y TDX.

Tesis doctorales

- Adaptación del modelo OAIS para su aplicación en hospitales y la preservación de las Historias Clínicas Electrónicas (presentada).
- Sistema de preservación digital de datos cartográficos (en curso).
- Técnicas de análisis forense aplicado a la preservación digital (en curso).
- Repositorio institucional y sistema de preservación para una universidad mexicana (en curso).

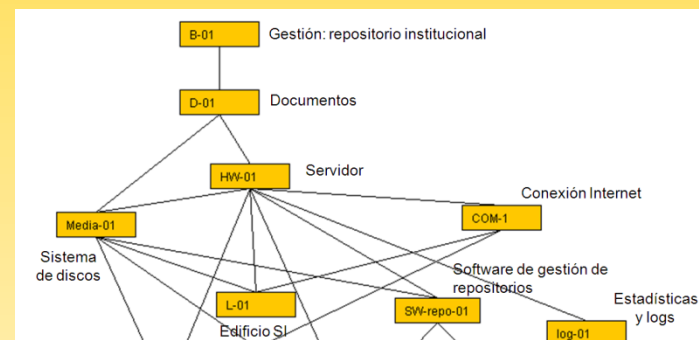
Consultoría

- Elaboración de esquemas METS y PREMIS.
- Integración y exportación de metadatos en repositorios universitarios y archivos digitales.
- Diseño y puesta en marcha de un sistema de preservación digital para la Universidad de Barcelona.

3. Resultados

Recomendaciones - 1

- Utilización de las auditorías de seguridad informática y de preservación digital como metodologías para la autoevaluación de la situación y de las perspectivas de mejora de una institución.
- Auditorías como herramienta de diagnóstico.
- Auditorías como herramientas de planificación.
- Auditorías como herramientas para conseguir la implicación de todos los actores.



3.3. [mp] Medidas de protección

control	[current]	[ENS]
[mp] Medidas de protección		L2
[mp.if] Protección de las instalaciones e infraestructuras		L2
[mp.if.1] Áreas separadas y con control de acceso		L2
[mp.if.2] Identificación de las personas		L2
[mp.if.3] Acondicionamiento de los locales		L2
[mp.if.4] Energía eléctrica		L2
[mp.if.5] Protección frente a incendios		L2
[mp.if.7] Registro de entrada y salida de equipamiento		L2
[mp.per] Gestión del personal		L2
[mp.per.2] Deberes y obligaciones		L2
[mp.per.3] Concienciación		L2
[mp.per.4] Formación		L2
[mp.eq] Protección de los equipos		L2
[mp.eq.1] Puesto de trabajo despejado		L2
[mp.com] Protección de las comunicaciones		L2
[mp.com.3] Protección de la autenticidad y de la integridad		L2
[mp.si] Protección de los soportes de información		L2
[mp.si.1] Etiquetado		L2
[mp.si.3] Custodia		L2

4.2.4.1.1	<i>The repository shall have unique identifiers</i>
3.1.2.1	<i>The repository shall have an appropriate succession plan, contingency plans, and/or escrow arrangements in place in case the repository ceases to operate or the governing or funding institution substantially changes its scope</i>



Recomendaciones - 2

- Adaptación de las técnicas del análisis forense digital para agilizar la ingestión de grandes volúmenes de datos.
 - Agrupación de datos en discos virtuales.
 - Descripción simplificada por lotes.
 - Acceso por emulación.

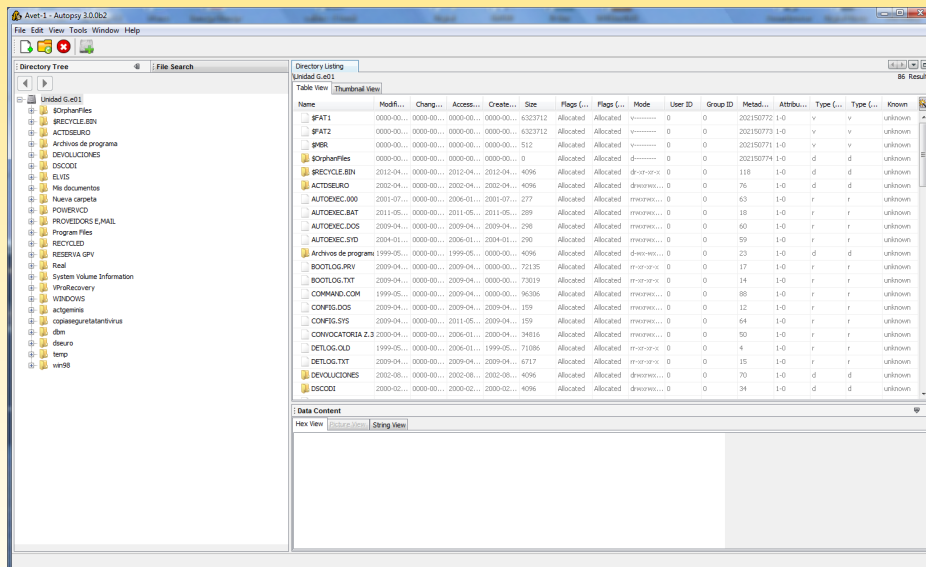


Imagen de un disco duro capturada con Autopsy



Ordenador de adquisición forense en la UB

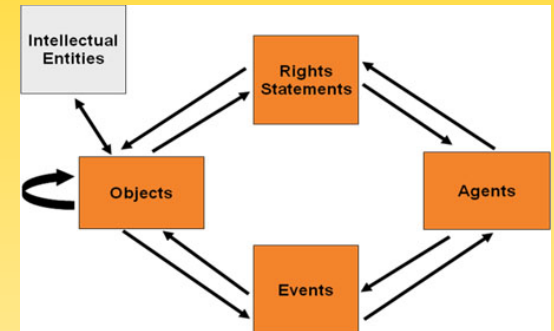
Recomendaciones - 3

- Aprovechamiento de los servicios en la nube (*cloud storage*) para dar seguridad al almacenamiento.
 - Actualmente el almacenamiento en la nube tiene un buen coste a corto plazo pero no a largo plazo.
 - Utilizarlo como solución para arrancar nuevos sistemas de preservación.
 - Buena solución para segundas copias (ej.: Amazon Glacier).



Recomendaciones - 4

- Revisión del uso dado al esquema de metadatos PREMIS.
- No obsesionarse con los esquemas XML; se pueden obtener resultados similares con bases de datos sencillas.
- Considerar PREMIS como una guía, no como un mandato.
- Por ejemplo, es más importante saber generar valores MD5 o SHA-1.



```
- <mets:amdSec ID="amd_for_id1">
-   <mets:techMD ID="tech_id1">
-       <mets:mdWrap MDTYPE="PREMIS">
-           <mets:xmlData>
-               <premis:object xsi:type="premis:file">
-                   <premis:objectIdentifier>
-                       <premis:objectIdentifierType>URL</premis:objectIdentifierType>
-                       <premis:objectIdentifierValue>/CTCRL11368/CONSULTA/CTCRL11368_001.jpg</premis:objectIdentifierValue>
-                   </premis:objectIdentifier>
-                   <premis:objectCharacteristics>
-                       <premis:compositionLevel>0</premis:compositionLevel>
-                       <premis:fixity>
-                           <premis:messageDigestAlgorithm>MD5</premis:messageDigestAlgorithm>
-                           <premis:messageDigest>d4b8e7444c586ba73b58b594796014a5</premis:messageDigest>
-                       </premis:fixity>
-                       <premis:size>1614885</premis:size>
-                   </premis:objectCharacteristics>
-               </premis:object>
-           </mets:xmlData>
-       </mets:mdWrap>
-   </mets:techMD>
</mets:amdSec>
- <mets:amdSec ID="amd_for_id2">
-   <mets:techMD ID="tech_id2">
```


Recomendaciones - 5

- Valoración de las soluciones de virtualización como una alternativa para el acceso a datos obsoletos.
 - Acceso por máquinas virtuales (ej.: VirtualBox, VMWare).
 - Uso de visores de formatos obsoletos (ej.: QuickView).



4. Conclusiones

Conclusiones

- Se han de conseguir soluciones eficaces, no perfectas.
- Existen muchas soluciones aplicables, de coste moderado y baja complejidad.
- Las soluciones presentadas no invalidan el marco teórico aportado por el modelo OAIS, lo hacen más fácilmente aplicable.

¡Muchas gracias!

Miquel Térmens Graells

Universidad de Barcelona

Departamento de Biblioteconomía y Documentación

termens@ub.edu

